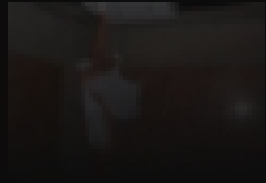
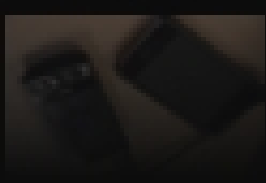
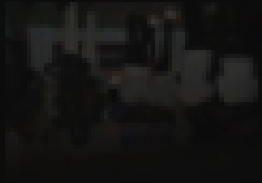
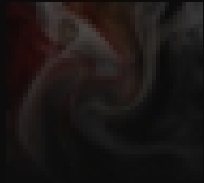
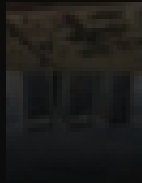
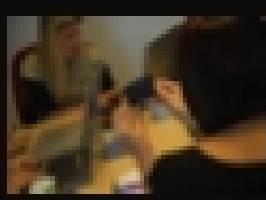
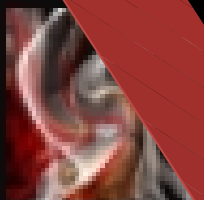
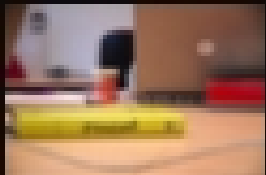
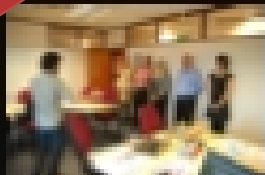
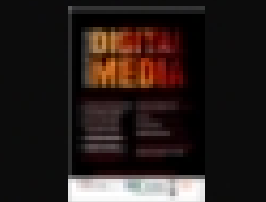
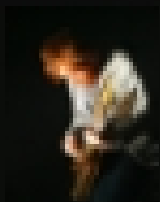
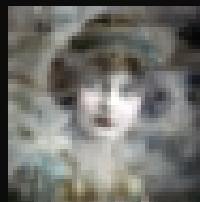
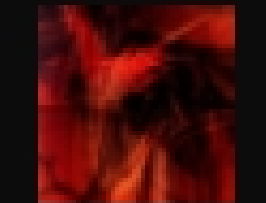
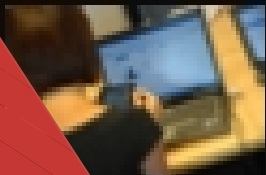
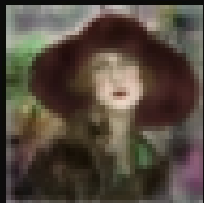
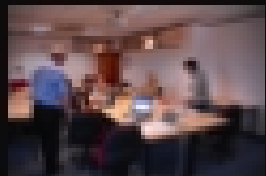
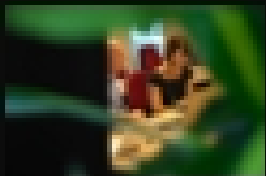
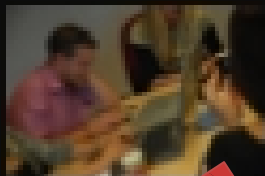
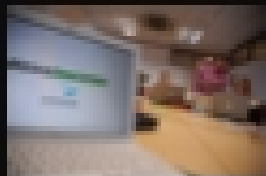
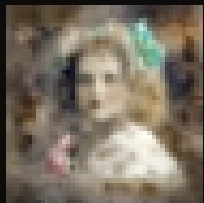


REACTION

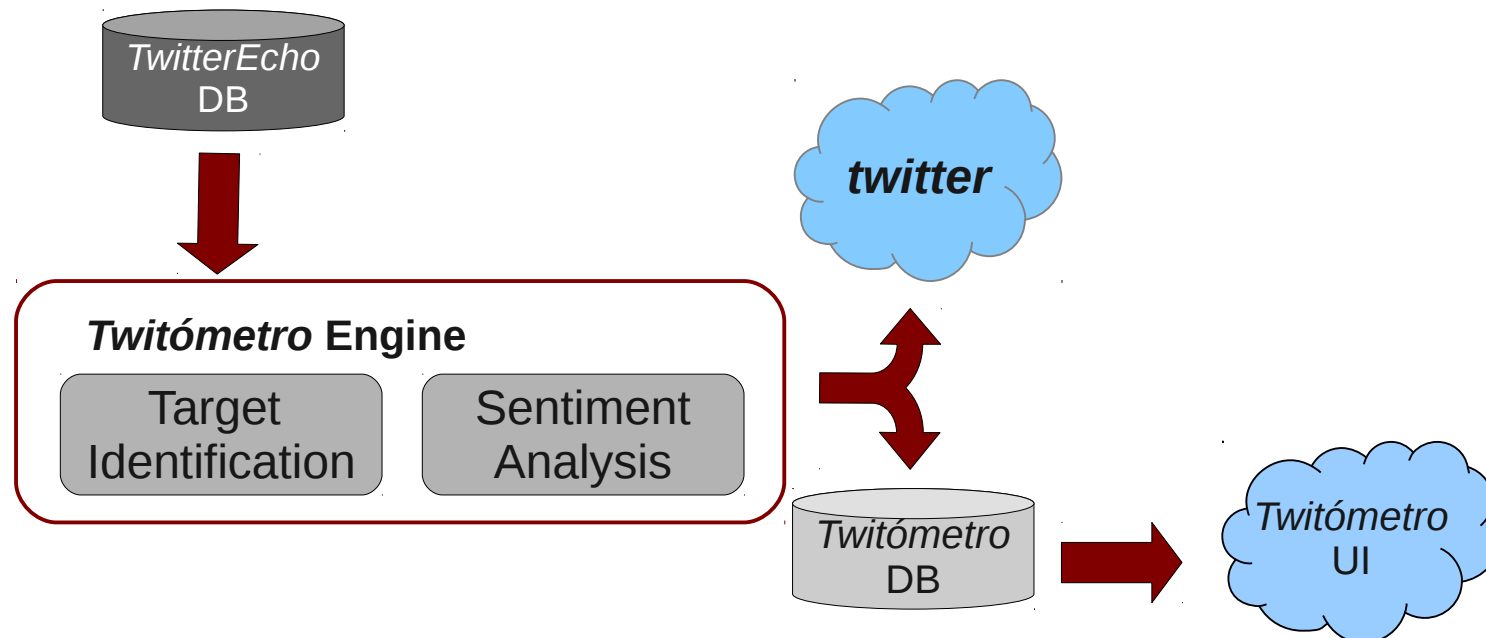


Summary

- 1. Process Overview**
- 2. Target Identification**
- 3. Lexical Resources**
- 4. Sentiment Analysis**
- 5. Results**
- 6. Next Steps**

Twitómetro Process Overview

1. Retrieve tweets from *twitterEcho*
2. Identify targets
 1. Discard tweets that don't refer any of the targets of interest
3. Perform sentiment analysis
4. Post results

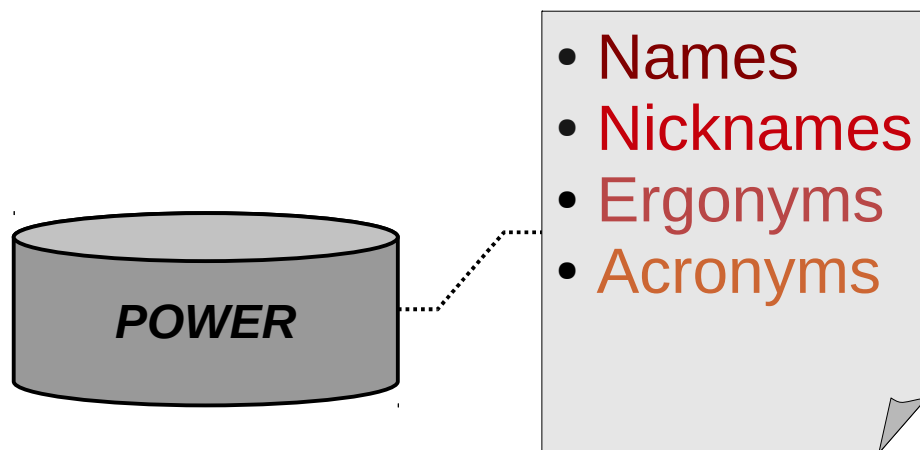


Twitómetro Target Identification

- Specialized dictionary with forms of mention for each target
- Set of exceptions for ambiguous words
 - “... *cinco portas* ...”
 - “... *portas abertas* ...”

- Ex.: Paulo Portas

- *paulo portas, portas*
- *paulinho das freiras, tio paulinho, dr. paulinho portas, paulinho das portas, paula das freiras*
- *candidato do cds/pp, candidato do cds, líder do cds-pp*
- *PP*



Twitómetro Lexical Resources

- Compilation of lexical resources and respective polarities
- *SentiTokens*
 - 6319 Adjectives
 - 966 Predicative nouns
 - 515 Verbs
 - 642 Idiomatic expressions
- Lemmas and inflected forms

Twitómetro Sentiment Classification

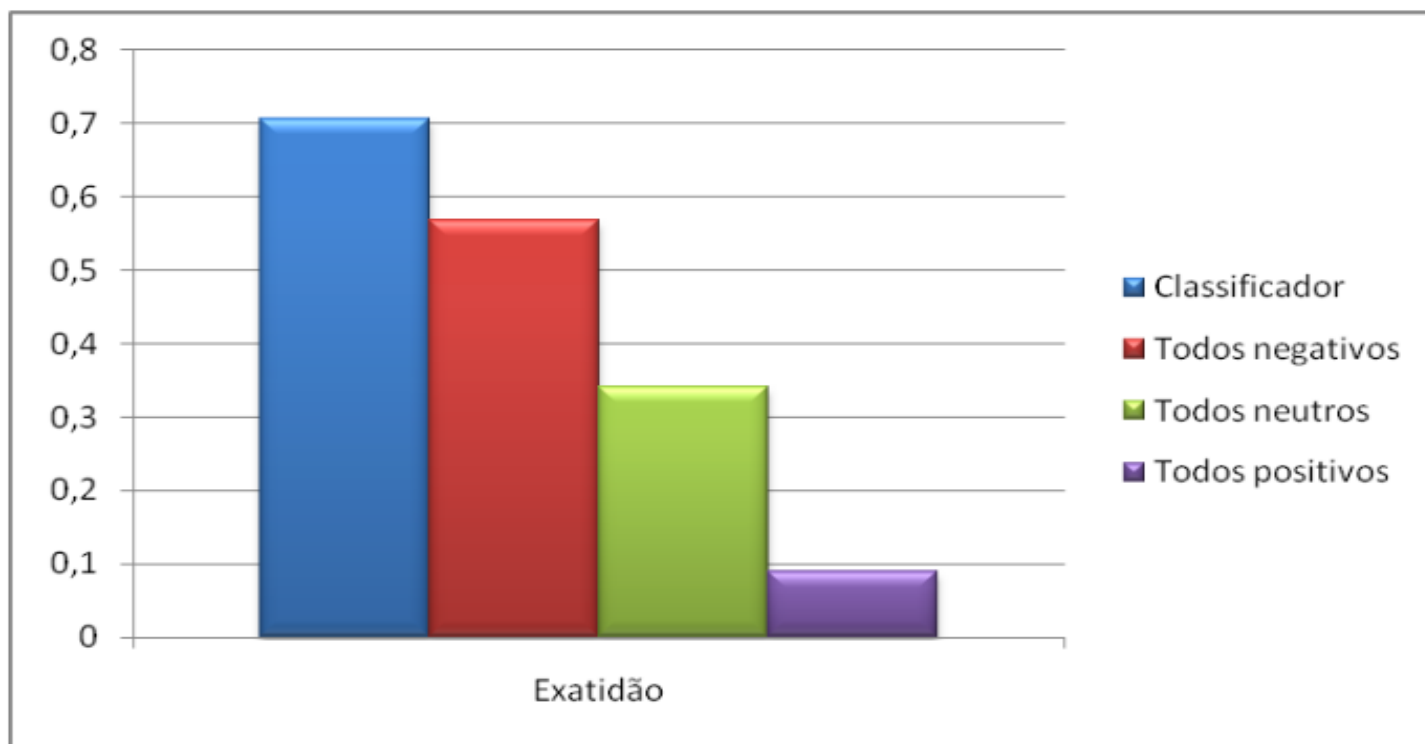
- Irony Detection
 - Quoted *sentiTokens* with prior positive polarity
 - *O Sócrates é um “excelente” ministro (Sócrates is an “excellent” minister)*
 - Mentions by nickname
- Emoticons :**)** :**P** :**D**
- Internet Slang *Lololol Hahahaha*
- Strong punctuation **!!!** **!?!?! ???**
- Interjections
 - *Oh meu deus! (Oh my god!)*

Twitómetro Sentiment Classification

- Rule-based engine
- Lexico-syntactic patterns
 - [copV] um [ADJ+] [ADJ-] → -1
 - ... *é um perfeito idiota* ... (is a perfect idiot)
 - não [copV] [ADJ-] → 1
 - ... *não é mentiroso* ... (is not a liar)
 - [copV] um [ADJ+] [N-] → -1
 - ... *é um verdadeiro desastre* ... (is a total wreck)
- Rules are ordered with a “longest path” criteria
- If none of the rules match:
 - Sum the polarities of all *sentiTokens*

Twitómetro Results

- Accuracy of target identification: **~99%**
- Accuracy of sentiment classification: **~71%**
- Contradiction rate: **~10%**



Twitómetro Next Steps

- Publish the corpus of analyzed tweets
- Improve sentiment classification with machine learning techniques
 - Use lexical evidences as features
- Expand lexical resources
- Expand lexico-syntatic pattern set
 - Include transitive verbs